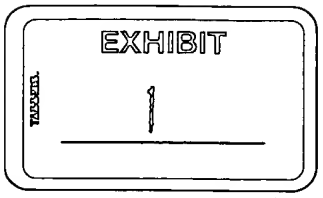
ANALYTICAL BIOCHEMISTRY 187, 364–373 (1990)




# Low-Ratio Hybridization Subtraction

Joseph Fargnoli,*[,1] Nikki J. Holbrook,* and Albert J. Fornace, Jr.†
*Laboratory of Molecular Genetics, NIA, NIH, Baltimore, Maryland 21224, and †Laboratory of Molecular Pharmacology, NCI, NIH, Bethesda, Maryland 20892

Received July 24, 1989

**A hybridization subtraction protocol that uses low ratios of RNA to cDNA has been developed to enrich for the cDNA of transcripts that are elevated in one cell population relative to another. This low-ratio hybridization subtraction protocol was found to yield substantial enrichment for the cDNA of low-abundance transcripts induced or increased only several fold. Conditions for the cloning of cDNA enriched by our hybridization subtraction and identification of clones coding for induced transcripts are presented. By screening the cDNA library with probes synthesized from the starting cDNA and cDNA enriched by low-ratio hybridization subtraction, clones coding for induced transcripts could be efficiently identified. The choice of reverse transcriptase used to synthesize the cDNA was found to be important for the enrichment of cDNA for longer length RNA. Low-ratio hybridization subtraction of cDNA synthesized with MMLV reverse transcriptase was effective for the enrichment of cDNA coding for RNA to at least 5 kb in length, while the AMV enzyme was effective only for the cDNA of shorter RNA (<1 kb). The characterization of several different low-ratio hybridization subtraction libraries is presented, and the advantages and disadvantages of various hybridization subtraction strategies are discussed.** 

cDNA clones of particular interest can be identified on the basis of their pattern of expression or relative abundance in different cells. The conventional approach has consisted of differential hybridization with labeled cDNA probes synthesized from RNA of either control cells or the experimental cells of interest (1). A cDNA clone corresponding to an RNA that is more highly expressed in the experimental cells relative to the control population will show a stronger hybridization signal with radiolabeled cDNA synthesized from the experimental RNA. However, in practical terms, when cDNA libraries are screened by differential hybridization, the lower limit of detection for a particular sequence is dependent on an abundance level of approximately ≥0.1% both in the cDNA probe and in the RNA from which it was synthesized (1). Considering that there are on average 2 to 5 × $10^5$ total mRNA transcripts present in a typical mammalian cell and probably 10,000 to 20,000 different mRNA species per cell (2), there would have to be at least 100 polyadenylated [poly(A)] transcripts per cell of a particular sequence in order to detect a cDNA clone by differential hybridization. Clearly, the vast majority of different transcripts in a mammalian cell are not this abundant and hence cannot be effectively detected by differential hybridization.

Enrichment of cDNA encoding preferentially expressed low-abundance (<0.05%) mRNA can be accomplished by hybridization subtraction (3–5). The basis of this approach is to enrich for sequences present in one RNA population, but absent, or at least present at much lower levels, in another RNA population by eliminating (subtracting) common sequences between the two populations. cDNA, which is synthesized from the experimental RNA, is hybridized in solution with a molar excess (10- to 60-fold) of control RNA to high $R_0t$ values in order to drive the hybridization to completion. Heteroduplex RNA/cDNA is separated from nonhybridized cDNA by techniques such as hydroxylapatite chromatography. The resulting cDNA, which is now enriched for sequences specific for the experimental cells, is then cloned according to standard procedures after second-strand synthesis (3).

To employ standard hybridization subtraction procedures, a specific sequence must be much more abundant in the experimental RNA sample relative to the control. For example, at a 25-fold excess of control RNA to cDNA, all cDNA encoding transcripts that are less than 25-fold more abundant in the experimental sample compared to the control should be removed by hybridization

[1] To whom correspondence should be addressed at Laboratory of Molecular Genetics, NIA, Gerontology Research Center, 4940 Eastern Avenue, Baltimore, MD 21224.



0003-2697/90 $3.00

subtraction. In theory, ratios of RNA:cDNA as low as 1:1 can be used if the hybridization reaction is driven toward completion by a high concentration of control RNA (driver); at such a low ratio, enrichment for sequences only several fold more abundant in the experimental RNA should be possible. Accordingly, we have developed a hybridization subtraction protocol that employs much lower ratios of control RNA:experimental cDNA than those used in standard hybridization subtraction protocols. This approach has enabled us to obtain substantial enrichment for sequences only several fold more abundant in the cDNA synthesized with the experimental RNA. Such an approach was necessary to isolate cDNA clones coding for DNA-damage-inducible (DDI)[2] transcripts in mammalian cells, because many DDI transcripts are of low abundance and are only induced several fold following treatment with DNA-damaging agents (6). In addition, to study complex and subtle cellular processes such as might occur during aging, relatively small changes in the expression of so-called "housekeeping genes" and other genes, whose transcript levels are of low abundance, might be expected. We have used hybridization subtraction at ratios of RNA:cDNA from 1:1 to 1:3 to construct five different cDNA libraries from mammalian cells and tissues. This report presents a detailed protocol for low-ratio hybridization subtraction and the optimization of this strategy as well as an overview of the libraries we have constructed using this approach.

## MATERIALS AND METHODS

A description of the methodology and preparation of reagents specifically relevant to low-ratio hybridization subtraction cDNA cloning has been emphasized here. The reader is referred to previously published reports for other techniques utilized for the library construction and screening (7–10).

*Solutions.* RNase-free precautions (1) were followed in the preparation of all solutions used in RNA preparation and analysis. Our precautions included the use of sterile glass-distilled water and HPLC-grade glass-filtered ethanol (Aldrich) and the use of freshly opened containers of salts or other solutes; the prepared solutions were filtered through two 0.22-$\mu$m Millex-GS sterile filters (Millipore) to bind and remove trace nucleases and other proteins.

All solutions used in hybridization subtraction were mixed with Chelex 100 (Bio-Rad) to remove divalent cations such as Mg or Fe; the Chelex 100 resin was removed with Millipore filtration or by centrifugation. Chelex 100 was stored in 1 M NaCl at 4°C and rinsed with water before use. Although probably unnecessary, when Centricon-30 microconcentrators (Amicon) were used (see below), both the RNA and the cDNA used in low-ratio hybridization subtraction were treated with Chelex 100; the resin was subsequently removed by centrifugation.

First-strand synthesis utilizing MMLV reverse transcriptase (BRL) was carried out in 60 mM KCl, 50 mM Tris (pH 7.6), 7 mM $MgCl_2$, 300 $\mu g \cdot ml^{-1}$ RNasin (Promega), 2 mM dithiothreitol, 60 $\mu g \cdot ml^{-1}$ actinomycin D, 7 $\mu g \cdot ml^{-1}$ oligo$(dT)_{12-18}$ (Pharmacia), 200 $\mu g \cdot ml^{-1}$ pd$(N)_6$ (Pharmacia), 0.75 mM dCTP, 50 $\mu Ci \cdot ml^{-1}$ [$^{32}$P]dCTP, 1 mM each dATP, dGTP, and dTTP, 40 $\mu g \cdot ml^{-1}$ poly(A)RNA, and 16,000 $\mu \cdot ml^{-1}$ MMLV reverse transcriptase (BRL). Second-strand synthesis was in 50 mM Hepes (pH 6.6), 5 mM Tris (pH 7.6), 6 mM $MgCl_2$, 5 mM $NH_4SO_4$, 100 mM KCl, 40 $\mu$M dNTPs, 100 $\mu Ci \cdot ml^{-1}$ [$^{32}$P]dCTP, 5 $\mu g \cdot ml^{-1}$ random 6-mers (PL), 0.1 mM dithiothreitol, 150 $\mu g \cdot ml^{-1}$ DNA polymerase (endonuclease-free, Boehringer Mannheim), and 50 $\mu g \cdot ml^{-1}$ BSA. cDNA tailing was in 0.2 M cacodylic acid, 25 mM Tris (pH 7.0), 80 mM KCl, 0.5 mM $CoCl_2$, 1 mM dCTP, 0.25 mg $\cdot ml^{-1}$ BSA, 2 mM dithiothreitol (added last; discard if solution turns brown). Annealing solution consisted of 100 mM NaCl, 10 mM Tris (pH 7.6), 1.0 mM EDTA (pH 7.5). All other solutions are described, where appropriate, in the remainder of the text.

*Cell lines and tissues.* Chinese hamster cell lines, lung (V79) fibroblast and ovary (CHO-K1), and human glioma cell lines, T98, A673, and U87MG (11), were grown as previously described (12,13). All experiments were performed with cells in the exponential growth phase.

Ten 5-month-old (young) and ten 24-month-old (old) outbred male Wistar rats, born and reared in the GRC aging animal colony (Baltimore), were used as the source for brain tissue. Human liver tissue was obtained from partial hepatectomy specimens of patients with metastatic tumor through the NCI Laboratory of Pathology and the NCI Surgery Branch; only normal-appearing liver tissue without tumor was used.

*cDNA clones and probes.* The Chinese hamster cDNA clones used were: pHS18A (12,16), which contained a 0.3-kb insert coding for B2 RNA; pH8 (16), which contained an 0.8-kb insert coding for hsp27 mRNA; pA2 (16), which contained a 1.1-kb insert coding for $\beta$-actin RNA; and p35 (13), which hybridized with a 2.5-kb poly(A)RNA. Under the heat shock (HS) conditions employed, B2 RNA was induced 5- to 10-fold, hsp27 RNA was induced 3.3-fold, and the level of RNA encoded by p35 remained constant (12). Table 1 summarizes the cDNA clones coding for uv-inducible transcripts that were isolated by low-ratio hybridization subtraction.

[2] Abbreviations used: DDI, DNA-damage-inducible; BSA, bovine serum albumin; HS, heat shock; TES, 2-[(tris(hydroxymethyl)-methyl)amino]ethanesulfonic acid; SDS, sodium dodecyl sulfate; HAP, hydroxylapatite.

cDNA that had been excised from its plasmid vector was labeled (14) and used at 3–12 × $10^6$ dpm·$ml^{-1}$. In order to prepare large numbers of different cDNA probes, cDNA was labeled in agarose by the random primer method (14) with minor modifications.

*RNA isolation and analysis.* RNA was isolated by solubilization in guanidine thiocyanate followed by ultracentrifugation on $CsCl_2$ cushions (17). It should be noted, however, that RNA isolated by a variety of other standard protocols is suitable for our approach. Poly(A) RNA was separated from total RNA by two cycles of oligo(dT) chromatography (1). RNA samples that were used as driver for hybridization subtraction were always treated with RNase-free DNase according to the manufacturer's suggestions (Worthington). The importance of this step is to eliminate any residual DNA that might ultimately be cloned in the final reaction steps, as we have detected trace amounts of DNA even after oligo(dT) chromatography when the DNAse step was omitted. Following DNAse treatment, the poly(A)RNA was extracted with chloroform/phenol (pH 7) and chloroform/isoamyl alcohol. The organic phases were back-extracted with water to recover most of the RNA; the poly(A)RNA was alcohol-precipitated and suspended in water at a concentration of <1 mg·$ml^{-1}$.

A critical point for both the low-ratio hybridization subtraction and the quantitative analysis of transcripts, which differ in the experimental and control population by only several fold, is that the RNA samples should contain equivalent amounts of total mRNA. Therefore, we measured the poly(A) content of different RNA samples using a labeled polythymidylic acid probe (12), and we used Northern blots with our β-actin probe to monitor for degradation. For RNA dot blot analysis, eight dilutions of each RNA sample were bound to nylon filters; the first four were hybridized to the probe of interest and the second four were hybridized with the polythymidylic acid probe. Results were normalized using a computer program ("RNA Analysis") which has been described previously (19). For hybridization with cDNA probes, the solutions described by Church and Gilbert (15) were used at 65°C (6,12,16,18). With this approach, differences of 1.5-fold or more in the level of a particular RNA could be consistently detected between control and experimental samples.

*Low-ratio hybridization subtraction.* Our low-ratio hybridization subtraction procedure is outlined in Fig. 1. The first step involves the synthesis of cDNA from poly(A) RNA. cDNA was synthesized with MMLV reverse transcriptase (BRL) in reverse transcriptase buffer or with AMV reverse transcriptase as previously described (12). The protocol for cDNA synthesis with MMLV reverse transcriptase included (i) briefly heating the RNA to >70°C, (ii) adding all the reagents except for the enzyme and incubating at 37°C for 5 min, (iii) incubating at 25°C for 5 min, (iv) adding the enzyme and incubating at 37°C for 30 min, (v) briefly incubating the solution at 25°C for 3 min, and (vi) incubating at 37°C for 1 h. Labeled deoxyribonucleoside triphosphate was included in the reaction mixture such that the specific activity of the cDNA was approximately $10^5$ dpm·$\mu g^{-1}$ ($5 \times 10^4$ Cerenkov cpm·$\mu g^{-1}$). The reaction was stopped by adding EDTA to 20 mM. The RNA was hydrolyzed at 65°C for 30 min in 150 mM NaOH. The solution was then neutralized by the addition of 2 M Tris, pH 7.5, to a final concentration of 150 mM. Low-molecular-weight molecules, including oligo$(dT)_{12-18}$, were removed with a Centricon-30 microconcentrator previously washed with TES (10 mM Tris, pH 7.5; 1 mM EDTA; 0.05% SDS). The yield of cDNA was determined both by the $OD_{260}$ and by the incorporation of labeled deoxyribonucleoside triphosphate. A range of 40–70 μg of cDNA was synthesized from 100 μg of poly(A) RNA. Approximately 1 μg of the cDNA was stored at −20°C and the remainder was purified by organic extraction. The cDNA was washed extensively with TrES (10 mM Tricine, pH 6.8; 0.4 mM EDTA, pH 7.5; 0.1% SDS) in a Centricon-30 microconcentrator in a fixed-angle rotor (SS34) which had been previously rinsed with RNase-free TrES. For the hybridization subtraction, driver (control) Poly(A) RNA (e.g., 80 μg) was added to the cDNA in the Centricon-30 microconcentrator and was subsequently washed with an additional 1 ml of TrES. The solution was recovered in approximately 50 μl from the Centricon-30 microconcentrator. In most experiments 2% of this sample was removed and stored separately at −20°C until it was used in a mock hybridization that contained no driver RNA in order to monitor the extent of self-annealing. The cDNA/RNA solution was dried in a Savant Speed Vac concentrator at 40°C to a final volume of approximately 15–20 μl. Water and 3 M sodium phosphate (pH 6.8) were added to yield a final volume of 25 μl and a final concentration of 0.4 M sodium phosphate. The preparation of the RNA and cDNA has been given in detail to emphasize the importance of having pure nucleic acids that are totally in solution and not aggregated. For example, when the Centricon-30 microconcentrator steps are replaced by alcohol precipitations, residual salt in the pellet often makes it difficult to dissolve the RNA completely at high concentrations. Some authors have found that hybridization is driven further to completion if glass beads are included in the hybridization and the tube is rotated to allow mixing by the beads (20); in our opinion, this observation probably reflects to some extent that the nucleic acids were not completely dissolved initially and that the continual mixing helped drive the nucleic acids into solution. In addition, if the sodium phosphate is added prior to the concentration in the Centricon-30 microconcentrator or if the solution freezes during the evaporation step, the nucleic acids may aggregate and precipitate out of solution. In order to

drive the hybridization as near as possible to completion, solution hybridization was carried out to high $R_0t$ values. The hybridization solution was transferred to a siliconized glass capillary pipet, sealed, heated briefly to 100°C, and then incubated for 25 to 35 h at 65°C. The final RNA concentration in our experiments was 2.5–4.0 $\mu g \cdot \mu l^{-1}$, and the uncorrected $R_0t$ was approximately $10^3$ assuming 10.3 $R_0t \cdot h^{-1}$ per $\mu g \cdot \mu l^{-1}$ of RNA (19).

Following the hybridization, single-stranded cDNA was separated by hydroxylapatite (HAP) chromatography at 62°C in the HAP solution (0.12 M sodium phosphate; 1 mM EDTA, pH 7.5; 0.2% SDS). Duplex RNA·cDNA and RNA will bind to the HAP at this temperature and phosphate concentration, while the single-stranded cDNA will elute through with several washes of the HAP solution. The duplex RNA·cDNA can be removed with >0.4 M phosphate or by heating the column to >90°C. As discussed by Britten *et al.* (19), successful HAP chromatography can be both an art and a science. Several suggestions are offered: (i) Use a water-jacketed column where the water jacket extends both below and above the resin to ensure uniform temperature. We routinely use columns (Bio-Rad) where the inner column has an inner diameter of 1 cm. The resin is poured to a packed height of 2.5 cm. In addition, the column is modified at both ends with stainless-steel seals in order to withstand the high temperatures used in HAP chromatography. (ii) There may be some variation between HAP lots. To test a lot, prepare radiolabeled single-stranded and double-stranded DNA; add the single-stranded DNA to the column and collect the single-stranded fraction; when this single-stranded fraction is added to a second column, essentially all of it should elute as single-stranded DNA. Do the same control with the radiolabeled double-stranded DNA; that is, all of the double-stranded fraction should elute as double-stranded DNA in a second HAP column. (iii) We use the standard grade HAP resin supplied by Bio-Rad. (iv) With certain lots (particularly very old ones), some high-molecular-weight HAP will solubilize and elute in the double-stranded fraction; this material is difficult to remove, precipitates in alcohol, and may interfere with the solubilization of trace quantities of DNA after alcohol precipitation. (v) Recently, some investigators have substituted biotinylated RNA in place of HAP chromatography. This approach, however, is not without its own problems; for example, at high hybridization temperatures, cDNA may be degraded in the presence of photobiotinylated RNA (data not shown).

To prepare for the second hybridization, the single-stranded cDNA was washed with TrES in a Centricon-30 microconcentrator to remove the phosphate. An equal amount of the same RNA as that used to synthesize the cDNA was added and washed once with TrES. The solution was concentrated as before and adjusted to 0.3 M sodium phosphate. Hybridization was carried out for 20–25 h at 65°C to a $R_0t$ of $>10^2$ in order to remove A + T-rich cDNA and other nonhybridizing species (6,12). The duplex RNA·cDNA fraction of the HAP chromatography was collected and the phosphate buffer removed in a Centricon-30 microconcentrator by washing with TES. It should be noted that in the presence of >0.05% SDS the recovery of single-stranded and double-stranded DNA from Centricon-30 microconcentrators was consistently >95% even with nanogram quantities. The solution was adjusted to 0.1 M NaOH, heated to 65°C, and neutralized with 2 M Tris, pH 7.5, as described and the hydrolyzed RNA was removed in a Centricon-30 microconcentrator by washing with TES. To remove the SDS, the DNA was alcohol-precipitated and resuspended in water. A small fraction of the enriched (prior to the second-strand synthesis step) cDNA (~10%) was saved for generating high-specific-activity enriched probe to be used in the library screening.

*cDNA cloning.* To synthesize double-stranded cDNA, enriched single-stranded cDNA was made double-stranded with DNA polymerase I and random primers under slightly acidic conditions to reduce exonuclease activity (14); the pH was then raised to reduce the frequency of gaps in the second strand. Single-stranded cDNA (100 ng) was briefly boiled and second-strand synthesis was carried out in 100 $\mu$l of the DNA polymerase I second-strand solution (6); the reaction was run for 30 min at 15°C and 15 min at 25°C. The pH of the solution was raised by the addition of 2 M Tris, pH 7.5, to a final concentration of 100 mM Tris, an additional 15 $\mu$l of DNA polymerase I was added, and the reaction was continued for 45 min at 25°C. Synthesis of second-strand cDNA was monitored by the incorporation of radiolabeled deoxynucleoside triphosphate. The yield of second-strand cDNA was routinely 50–70 ng. The reaction was stopped with 15 mM EDTA, and protein was removed by organic extraction. The double-stranded cDNA was size-separated on Sepharose-4B (Pharmacia). cDNA of 250 bp or more was selected and concentrated in a Centricon-30 microconcentrator followed by two washes with 2 ml of TES and a final wash in $H_2O$. In some experiments the DNA was alcohol-precipitated, while in others the sample was used directly for tailing.

The cDNA was cloned following standard GC tailing (1). Optimization of tailing was calibrated based on the optimal transformation efficiency as described previously (21). Briefly, small portions of the cDNA tailing reaction were removed at 0.5, 1, 2, and 3 min; the reaction was stopped by adding EDTA to a final concentration of 10 mM and heating to 65°C for 5 min. These samples were then annealed to the G-tailed vector in varying ratios of vector:cDNA in the annealing buffer. A typical annealing reaction included 36 ng of vector and 2 ng of cDNA in a 50-$\mu$l volume; the sample was placed in a 65°C

water bath where the thermostat was adjusted to 58°C for 1 h, and then placed in a 56°C bath for 2 h. A volume of 5 μl of the annealing buffer was added to 100 μl of transformation-competent DH5 *Escherichia coli* bacteria (BRL). Transformation was carried out according to the supplier's recommendations. Routinely, 1000–5000 colonies were obtained per nanogram of double-stranded cDNA (per 18 ng of vector DNA). With DH5 bacteria, G-tailed vector alone resulted in approximately 20 colonies per nanogram; this can be a problem when the cDNA clones poorly. An alternative is to use transformation-competent HB101 *E. coli* (BRL); with this strain, the transformation efficiency was nearly as high as that for DH5, but the frequency of background colonies with vector alone was reduced 10-fold (data not shown).

*Colony-blot hybridization.* As outlined in Fig. 1, cDNA clones coding for transcripts elevated in the experimental relative to control cells were tentatively identified by preferential hybridization with radiolabeled cDNA enriched by subtraction (enriched cDNA). Therefore, there are two requirements in the choice of vector used; they are the ease of replicate plating and a strong signal for colony hybridization. Our libraries were constructed in the pXF3A plasmid vector (6), a derivative of pXF3 (22) lacking the 0.55-kb *Pvu*II–*Aat*II region. This vector resulted in stronger signals by colony-blot hybridization than pBR322 or pUC plasmids when amplified with chloramphenicol; however, it lacked convenient restriction enzyme sites. One might consider substituting another high-copy-number plasmid vector (such as pUC which does not require chloramphenicol amplification) with more convenient restriction enzyme sites. To store the bacteria in a convenient form for replicate plating, colonies were picked with sterile toothpicks from the original plates, transferred to 96-well microtiter plates containing freezing medium (3.15 g $K_2HPO_4$, 0.9 g $KH_2PO_4$, 0.225 g sodium citrate, 0.045 g $MgSO_4 \cdot 7H_2O$, 0.45 g $(NH_4)_2SO_4$, 22 ml glycerol, and 6.25 mg tetracycline·HCl added to 500 ml standard Luria-Bertani [LB] broth), grown overnight at 37°C, and then stored frozen. For replicate plating, the bacteria (3000–4000 colonies) were thawed and transferred with a 48-well metal applicator (applicator tips should have a diameter of 2 mm for optimal colony size; commercially available from Watson Products, Natick, MA) to 36 × 24-cm sheets at approximately 1000 colonies per sheet of nitrocellulose or nylon filters which had been laid over 38 × 26-cm stainless-steel trays containing LB agar with antibiotic. The bacteria were incubated overnight at 37°C, and then the filters were transferred to trays containing LB agar with 170 $\mu g \cdot ml^{-1}$ chloramphenicol for an additional 24 h. The bacteria were lysed and bound to the filters by standard techniques (1). The filters were prehybridized 6–15 h in blot buffer (1 M NaCl; 0.2 M Tris, pH 7.2; 0.1% sodium pyrophosphate; 1 mM EDTA; 0.5% SDS; 1× Denhardt's solution [1]; 0.1 $mg \cdot ml^{-1}$ single-stranded salmon sperm DNA; 0.1 $mg \cdot ml^{-1}$ yeast tRNA; 0.1 $mg \cdot ml^{-1}$ polyadenylic acid) for nitrocellulose filters or Church hybridization buffer (15) for nylon filters, hybridized overnight at 65°C with $10^6$ $dpm \cdot ml^{-1}$ of labeled cDNA in large bags such that there were only three filters per bag, and washed the next day at 65°C in library wash buffer (1× SSC [1], 0.5% SDS, 0.1% sodium pyrophosphate, 1 mM EDTA) for nitrocellulose or wash buffers A and B (15) for nylon filters. Colonies having higher signals with the enriched cDNA probe relative to unenriched cDNA probe were selected for further characterization.

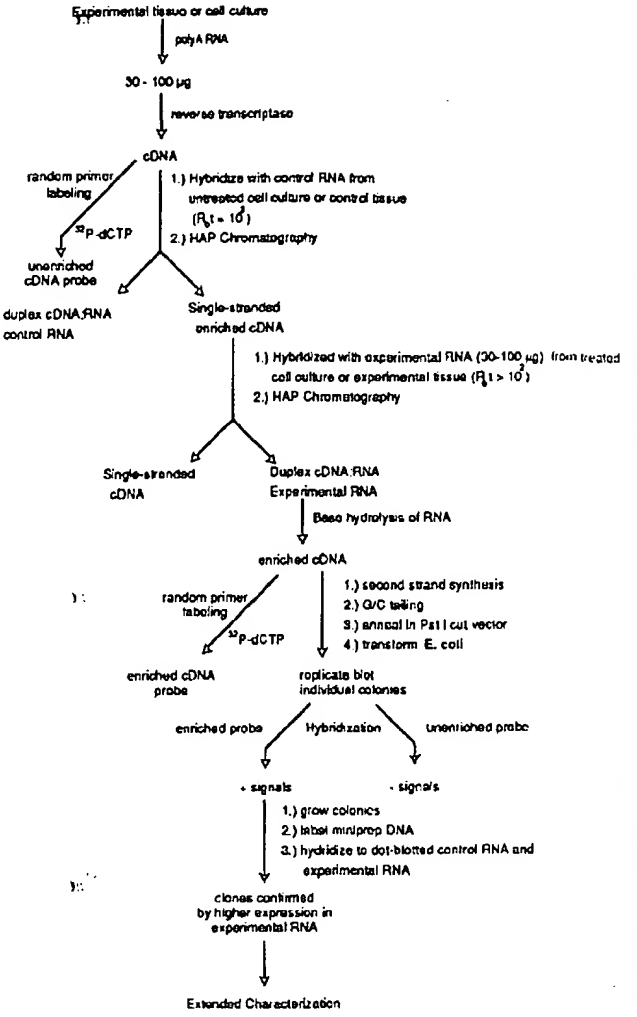



**FIG. 1.** Low-ratio hybridization subtraction and cDNA cloning scheme.

## RESULTS AND DISCUSSION

The procedure for low-ratio hybridization subtraction and cDNA cloning has been outlined in Fig. 1. Library

TABLE 1

DNA-Damage-Inducible cDNA Clones Isolated by Hybridization Subtraction[a]

| cDNA clones | Number of isolates[b] | RNA (kb)[c] | Induction[d] | cDNA clones | Number of isolates[b] | RNA (kb)[c] | Induction[d] |
|---|---|---|---|---|---|---|---|
| CHO cell library (library A) | | | | CHO cell library (library A) (*continued*) | | | |
| A29 | 2 | 4.8 | 3 | A33 | 1 | 0.92 | 1.9 |
| A141 | 1 | 4.3 | 3.1 | A7 | 1 | 0.8 | 2 |
| A31 | 2 | 3.9 | 2.1 | A50 | 1 | 0.8 | 2.4 |
| A99[e] | 1 | 3.9 | 2 | A87 | 1 | 0.67 | 1.8 |
| A143 | 1 | 3.7 | 4 | A162 | 1 | 0.6 | 2 |
| A88 | 1 | 3.7 | 1.6 | A115 | 4 | 0.4 | 2 |
| A8 | 1 | 3.1 | 3.5 | A72 | 1 | 0.35 | 2 |
| A112 | 1 | 2.7 | 1.8 | A83 | 1 | | 1.6 |
| A170[e] | 1 | 2.7 | 1.9 | A180 | 1 | | 3.9 |
| A78 | 1 | 2.4 | 1.6 | A94 | 1 | NV[f] | 4.1 |
| A15 | 1 | 2.4 | 1.6 | A70 | 2 | NV | 2.4 |
| A34 | 1 | 2.2 | 2.3 | V79 cell library (library U) | | | |
| A84 | 1 | 2.1 | 2.9 | U1 (metallothionein II) | 14 | 0.54 | 28 |
| A20 | 2 | 2.0 | 4.9 | U2 (metallothionein I) | 13 | 0.51 | 6.0 |
| A9 | 8 | 1.9 | 5.3 | U4 | 3 | 0.47 | 2.0 |
| A13[e] | 1 | 1.9 | 3.4 | U49 | 1 | 0.73 | 4.0 |
| A148 | 1 | 1.9 | 4.1 | U56 | 1 | 0.92 | 3.5 |
| A71 | 1 | 1.75 | 1.6 | U52 | 2 | 0.97, 0.68 | 4.6 |
| A109 | 1 | 1.4 | 3.9 | U53 | 1 | 1.27 | 2.2 |
| A18 | 4 | 1.4 | 3.1 | U5 | 3 | NV[f] | 15 |
| A45 | 1 | 1.4 | 4.2 | U64 | 1 | NV | 2.4 |
| A77 | 2 | 1.37 | 2.5 | U30 | 1 | NV | 2.5 |
| A26 | 2 | 1.3 | 2.5 | U44 | 1 | | 2.1 |
| A153 | 5 | 1.0 | 3.1 | U45 | 1 | | 2.2 |
| A90 | 1 | 1.0 | 1.7 | Actin | | | 0.9[g], 0.94[g] |
| A185 | 1 | 1, 0.77 | 1.6 | | | | |

[a] cDNA clones, coding for DNA-damage-inducible transcripts, were isolated from Chinese hamster cells. Library A was derived from Chinese hamster ovary cells harvested 4 h after 14 J·m$^{-2}$ of far-uv irradiation and library U from V79 cells treated in a similar manner (6). cDNA was synthesized with MMLV reverse transcriptase for library A and with AMV reverse transcriptase for library U.

[b] cDNA clones were considered to be derived from the same transcript when they cross-hybridized at high stringency and hybridized to the same transcript.

[c] The size of the transcript was determined by Northern blots.

[d] Relative induction was determined by RNA dot-blot hybridization using the same poly(A) RNA from uv-irradiated cells and untreated cells as that used in the low-ratio hybridization subtraction.

[e] While certain cDNA clones did not cross-hybridize, they appeared to code for different portions of the same transcript based on northern blots and their kinetics of RNA induction. Such clones included A31 vs A99, A112 vs A170, and A9 vs A13.

[f] Northern blots with certain cDNA probes resulted in faint broad smears rather than hybridization to specific bands and have been designated NV (not visualized). By Southern blots, U5 appeared to be repetitive DNA while A94 hybridized to a single locus.

[g] Relative β-actin mRNA in uv-irradiated Chinese hamster ovary cells and V79 cells, respectively.

A, which was constructed with RNA from uv-irradiated CHO cells (Table 1), will be used as a model for an explanation of this approach. In this experiment 80 μg of poly(A) RNA was used to synthesize the cDNA that was hybridized with 80 μg of poly(A) RNA from control (untreated) cells. After this hybridization, 9.8% of the starting cDNA eluted in the single-stranded fraction from the HAP column. In this and all other experiments, we observed that a substantial portion of this single-stranded fraction consisted of cDNA that could not form stable duplexes under our conditions. For example, when this cDNA was hybridized with excess control RNA or even the experimental RNA that had been used to synthesize the cDNA, a substantial fraction always remained single stranded. For this reason, the single-stranded cDNA from the first hybridization was hybridized a second time to the experimental RNA in order to remove these nonhybridizing species; the identity of this nonhybridizing cDNA is unknown but probably includes very AT-rich sequences. The second hybridization was always carried out to a $R_0t$ high enough to ensure that even low-abundance sequences would hybridize; e.g., in a typical mammalian cell the $R_0t_{1/2}$ for even low-abundance mRNA is less than 50 (19). Even when the second hybridization was carried out to $R_0t$ values approaching $10^3$, a substantial fraction of the cDNA remained single stranded (data not shown). Thus, this nonhybridizing cDNA probably cannot form a stable duplex under our

conditions regardless of the $R_0t$ value. In all our experiments, an enrichment of 1.5- to 4-fold was obtained with this second hybridization. In the case of library A, 36% of the cDNA in the second hybridization eluted in the double-stranded fraction. Thus, the theoretical enrichment after the two hybridizations was calculated to be 29-fold on the basis that 3.5% of the starting cDNA remained after low-ratio hybridization subtraction. As described under Materials and Methods, second-strand cDNA was synthesized using random primers and the cDNA was cloned into a plasmid vector by GC-tailing. It should be pointed out that the average size on denaturing gels for the cDNA remaining after low-ratio hybridization subtraction was substantially shorter than that for the starting cDNA. This might be expected for a variety of reasons which will not be discussed here but would certainly include the extensive experimental manipulations that the cDNA undergoes prior to cloning. Because of their short size, the enriched cDNA was fractionated on Sepharose-4B and fragments greater than 250 bp in length were cloned.

*Identification of cDNA clones coding for induced RNA.* Following construction of an enriched cDNA library, several different approaches can be employed to identify clones of interest. The most direct would be to pick random clones and screen much of the library. As described earlier, cDNA encoding uninduced transcripts are reduced by low-ratio hybridization subtraction, but are not entirely "subtracted out"; thus, the direct approach of screening random clones would be very labor intensive, since many would not encode induced transcripts. Our approach has been to tentatively identify clones coding for induced transcripts on the basis of preferential hybridization to the enriched cDNA (see Fig. 1) compared to the unenriched cDNA (cDNA not fractionated by hybridization subtraction). This approach should be much more sensitive than standard differential hybridization using unenriched cDNA. As discussed earlier, the lower limit of detection by colony hybridization for a particular sequence is dependent on an abundance level of approximately 0.1% or more in the probe. Since low-ratio hybridization subtraction can produce an enrichment of 30-fold or more for the cDNA of induced transcripts (12), the abundance for such a transcript in poly(A) RNA could be 0.003% or less (<6 transcripts per cell) and still be detected by screening with the enriched cDNA. An example of this approach is shown in Fig. 2. As seen in panel B, many cDNA clones that showed little or no detectable hybridization (above background) to the unenriched cDNA probe showed obvious hybridization to the enriched cDNA probe. To confirm that the clones identified by colony hybridization coded for induced transcripts, probes were prepared from individual cDNA clones and hybridized to dot blots containing the same poly(A) RNA as that used in the

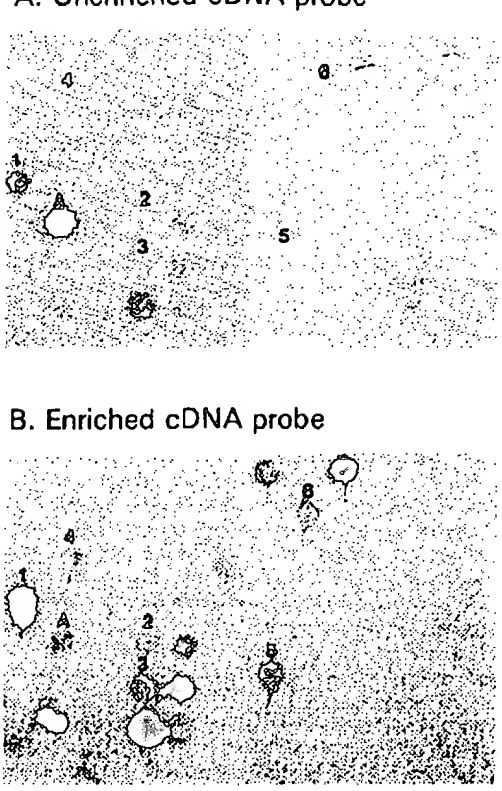



FIG. 2. Colony-blot hybridization. Replicate filters of cDNA library A (Chinese hamster ovary cells) were hybridized with either probe, which were synthesized with either the original (prior to hybridization subtraction) cDNA (A) or the cDNA enriched by low-ratio hybridization subtraction (B); see Materials and Methods for details. Numbers above a particular colony represent some of the clones that preferentially hybridized with the enriched cDNA; the colony designated by the letter A represents an abundant sequence that was not enriched, but was reduced (subtracted) by low-ratio hybridization subtraction.

low-ratio hybridization subtraction. As described under Materials and Methods, 20 or more cDNA probes could be prepared in 1 day; the hybridizations and washings were conveniently carried out in screw-top test tubes (such as Falcon No. 3033) such that 20 to 40 samples could be processed in 1 day. As shown in Table 2 60 to 70% of clones in libraries A, U, and L, which had been identified in this manner, were found to code for transcripts that were more abundant in the experimental than in the control RNA.

*Enrichment for cDNA synthesized with MMLV or AMV reverse transcriptase.* The choice of reverse transcriptase was found to be critical in our experiments. For example, in Table 1, the cDNA clones of library A were found to code for transcripts up to 5 kb in length, while the clones of library U only coded for short transcripts. The important difference appeared to be the use of the AMV reverse transcriptase. This was con-

TABLE 2

Comparison of Different Low-Ratio Hybridization Subtraction cDNA Libraries

| Library | Template for cDNA[a] | Driver RNA[b] | RNA:cDNA | Enriched clones[c] | Confirmed by RNA dot blots[d] |
|---|---|---|---|---|---|
| Library A[e] | CHO (uv) | CHO | 1:1 | 424 | 89/145 (61%)[f] |
| Library U[e] | V79 (uv) | V79 | 3:1 | 64 | 45 (70%) |
| Library L[g] | Liver | mer$^-$ | 2:1 | 182 | 109 (60%) |
| Young-enriched[h] | Brain (young) | Brain (old) | 1.5:1 | 74 | 20 (27%) |
| Old-enriched[h] | Brain (old) | Brain (young) | 1.5:1 | 87 | |

[a] RNA used as template for cDNA synthesis; see Materials and Methods. CHO (Chinese hamster ovary).

[b] Driver RNA of first hybridization subtraction (see Fig. 1) with the exception of library L (see below).

[c] cDNA clones that hybridized preferentially with the respective enriched cDNA; see Materials and Methods.

[d] cDNA clones that hybridized more strongly (≥twofold) with the RNA that had been used for cDNA synthesis compared to RNA that had been used for driver RNA.

[e] All the libraries with the exception of library U were constructed with cDNA synthesized with MMLV reverse transcriptase. See Table 1 for characterization of libraries A and U.

[f] Only a portion (145) of the cDNA clones, which preferentially hybridized with the enriched cDNA, were used for RNA dot blots.

[g] The protocol for the low-ratio hybridization subtraction of this library differed from Fig. 1. cDNA was first hybridized with RNA from T98 cells (mer$^+$) and the duplex cDNA was recovered; this cDNA was then hybridized with equal amounts of RNA from two mer$^-$ cell lines, A673 and U87MG, and the single-stranded cDNA was isolated.

[h] RNA was obtained from the brains of 5-month-old rats (young) and 24-month-old rats (old); see Materials and Methods.

firmed in the experiment shown in Fig. 3. Poly(A) RNA was isolated from heat-treated cells, cDNA was synthesized with either MMLV or AMV reverse transcriptase, and low-ratio hybridization subtraction was carried out as outlined in Fig. 1. Probes were synthesized with either the original or the enriched cDNA and hybridized to filters containing plasmid DNA coding for a short heat-inducible RNA (lane 1), a longer length heat-inducible RNA (lane 2), or a high-abundance RNA that was not heat-inducible (lane 3). As seen in panel B compared to panel A, when cDNA was synthesized with MMLV, there was substantial enrichment by low-ratio hybridization subtraction for both short and long heat-inducible sequences, while the uninduced sequence was clearly reduced (subtracted) after low-ratio hybridization subtraction. In contrast, when cDNA was synthesized with AMV reverse transcriptase (panels C and D), enrichment was only seen for the cDNA of the short heat-inducible RNA (lane 1) with no enrichment for the cDNA of the longer length heat-inducible RNA (lane 2). Enrichment only for the cDNA of short heat-inducible RNA was consistently found in a variety of other experiments when AMV reverse transcriptase was used (data not shown). This result may be due to the higher RNase H activity in AMV reverse transcriptase (compared to the MMLV enzyme) that results in nicking of the RNA template and synthesis of limited regions of second-strand cDNA, particularly for longer length sequences. At the high $C_0t$ values achieved during low-ratio hybridization subtraction, this second-strand cDNA could form duplex with the first-strand cDNA. This interpretation is supported by the results in Table 3 from experiments in which mock low-ratio hybridization subtraction was performed without driver RNA. At low $C_0t$ values both MMLV- and AMV-synthesized cDNA remained single stranded, which rules out the presence of rapidly reannealing "hair-pin" cDNA. However, at high $C_0t$ values a substantial fraction of the AMV reverse transcriptase cDNA formed duplex; this effect was not seen with several different lots of MMLV reverse transcriptase (provided by BRL). Results similar to those shown in Table 3 were obtained with 10 different lots of AMV reverse transcriptase provided by three different sources (BRL, Life Sciences, Seikagaku) (data not shown). While the endogenous RNase H activity in the AMV enzyme may be reduced with pyrophosphate or spermidine, inclusion of pyrophosphate with the AMV enzyme in an experiment similar to that summarized in Table 3 did not substantially reduce the fraction of second-strand cDNA in our hands (data not shown). Thus, the use of the MMLV reverse transcriptase is recommended for low-ratio hybridization subtraction particularly if the sequences of interest are derived from longer length transcripts. Recently, recombinant MMLV reverse transcriptase, which lacks RNase H activity, has become commercially available; while we have no experience with it, this new enzyme may be ideal for low-ratio hybridization subtraction.

*Characterization of low-ratio hybridization subtraction libraries.* Of the libraries listed in Table 2, the two enriched for DNA-damage-inducible sequences have been most extensively characterized and demonstrate some factors that should be considered when planning for the construction of a low-ratio hybridization subtraction library. In contrast to the other libraries of this table, the hybridization subtractions of libraries A and U em-

ployed RNA and cDNA from only one cell type each. The cells were treated with a DNA-damaging agent at a dose at which the primary cellular target should be only DNA. In addition, the cells were briefly incubated after irradiation in order to enrich for the cDNA of transcripts directly induced and to avoid changes in RNA levels that might occur at later times due to cell lethality and changes in cell cycle. In addition to the obvious advantage of using MMLV reverse transcriptase for library A, the cells, which this library was constructed with, do not express metallothionein RNA which are abundant uv-inducible transcripts in V79 cells; as seen in Table 1, many of the clones from library U were found to code for metallothionein. Following the isolation of these clones, the kinetics of induction and the spectrum of induction by different DNA-damaging agents or other stresses were studied (6); such studies are critical when trying to determine the relevance of changes in RNA levels of only several fold.

In contrast to libraries A and U, the construction of the other libraries of Table 2 had several disadvantages—in particular, the use of different cell types and higher RNA complexity. The strategy for library L was to enrich for sequences more abundant in mer$^+$ cells than mer$^-$ cells. Cells with the mer$^-$ phenotype have an increased sensitivity to certain alkylating agents and a reduced level of activity for $O^6$-alkylguanine–DNA alkyltransferase, a protein involved in the repair of DNA alkylation damage (11). cDNA was synthesized from the RNA of human liver which contains high levels of $O^6$-alkylguanine–DNA alkyltransferase. The cDNA was first hybridized with RNA from a mer$^+$ human glioma cell line and the common sequences (54% was duplex RNA:cDNA) recovered; this cDNA was then hybridized with RNA from two mer$^-$ glioma cell lines and the non-hybridizing cDNA (10% was single stranded) was recovered. As seen in Table 2, a large number of cDNA clones were isolated that coded for transcripts which were more abundant in the mer$^+$ glioma line than in the 2 mer$^-$ glioma lines. Our strategy was to determine the RNA levels for many of these cDNA clones in a panel of mer$^+$ and mer$^-$ cell lines. However, during the course of these studies, $O^6$-alkylguanine–DNA alkyltransferase clones were identified by another approach (manuscript in preparation); in addition, our approach using the panel of mer$^+$ and mer$^-$ cell lines was confounded by the substantial variability in the expression of this transcript in different mer$^+$ cell lines. In the last two libraries of Table 2, brain tissue from old and young animals was compared. The strategy was to enrich for "housekeeping" genes whose levels might be fairly constant in different cell types but change with aging. As seen in Table 2, few cDNA clones in the young-enriched library were identified that were more abundant in the RNA from young compared to old animals; when RNA from brains of a panel of different animals was screened with these cDNA probes, most were found not to be consistently elevated in younger animals. If such age-specific transcripts do exist, alternate strategies will probably have to be used. The high complexity of RNA from whole tissues and differences in the proportion of various cell types between animals in both young and old individuals were probably two serious disadvantages of our approach. Another concern, when planning low-ratio hybridization subtraction, is the number of different transcripts whose levels vary between the control and

TABLE 3

Comparison of cDNA Synthesized with AMV or MMLV Reverse Transcriptase[a]

| Enzyme | Yield[b] (%) | Duplex cDNA[c] (%) $C_0t < 10^{-1}$ | Duplex cDNA[c] (%) $C_0t > 10^2$ |
|---|---|---|---|
| AMV | 40–70 | <10 | 31 ± 22 |
| MMLV | 40–70 | <5 | 6.6 ± 3.2 |

[a] Data represent pooled results from different low-ratio hybridization subtraction experiments performed on different days with RNA isolated from Chinese hamster ovary cells.

[b] The yield was estimated based on the amount of cDNA synthesized compared to the amount of poly(A) RNA used as template.

[c] As described under Materials and Methods, a mock low-ratio hybridization subtraction was carried out in each experiment with cDNA only. Results represent the fraction of cDNA that formed duplex at either a low or high $C_0t$; values represent the mean with SD.

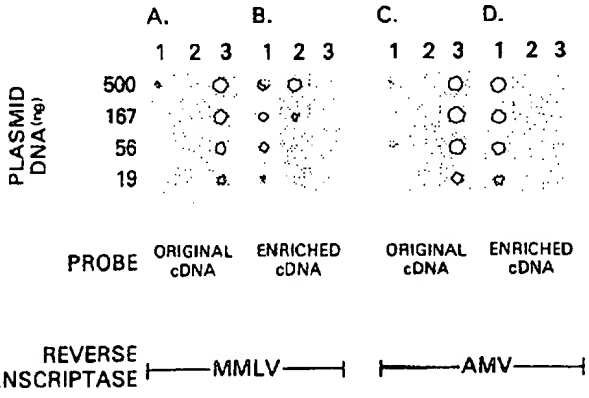



FIG. 3. Enrichment by low-ratio hybridization subtraction for cDNA synthesized with MMLV or AMV reverse transcriptase. Low-ratio hybridization subtraction was used to enrich for heat shock (HS) inducible cDNA at a 1:1 ratio of RNA:cDNA. Probes were then synthesized with either the original cDNA (prior to low-ratio hybridization subtraction) or the enriched cDNA and hybridized to replicate DNA dot blots: lane 1, pHS18A DNA which codes for a short transcript that was induced 5- to 10-fold by HS; lane 2, pH8 which codes for a longer length transcript that was induced 3.3-fold by HS; and lane 3, p35 which codes for a transcript not induced by HS (12). The autoradiograph exposure times were different for lanes 1, 2, and 3, but remained constant in panels A to D.

experimental cell types. For example, if a large number vary substantially, such as ≥5%, then low-ratio hybridization subtraction may not be effective; preliminary experiments, such as two-dimensional gel analysis of cellular proteins, should be considered in such cases.

## REFERENCES


1. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
2. Lewin, B. (1985) Genes II, p. 307, Wiley, New York.
3. Sargent, T. D., and Dawid, I. B. (1983) *Science* **222**, 135–139.
4. Hedrick, S. M., Cohen, D. I., Nielsen, E. A., and Davis, M. M. (1984) *Nature (London)* **308**, 149–153.
5. Timberlake, W. E. (1980) *Dev. Biol.* **78**, 497–510.
6. Fornace, A. J., Jr., Alamo, I., Jr., and Hollander, M. C. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 8800–8804.
7. Rhyner, T. A., Bignet, N. F., Berrard, S., Barbely, A. A., and Mellet, J. (1986) *J. Neurosci. Res.* **16**, 167–81.
8. Meruelo, D., and Rossomando, A. (1986) *Leuk. Res.* **10**, 819–832.
9. Kavathas, P., Sakhatme, V. P., Herzenberg, L. A., and Parnes, J. R. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 7688–7692.
10. Vitek, M. P., Kreissman, S. G., and Gross, R. H. (1981) *Nucleic Acids Res.* **11**, 1191–202.
11. Yarosh, D. B. (1985) *Mutat. Res.* **145**, 1–16.
12. Fornace, A. J., Jr., and Mitchell, J. B. (1986) *Nucleic Acids Res.* **14**, 5793–5811.
13. Fornace, A. J., Jr. (1983) *Nature (London)* **304**, 552–554.
14. Feinberg, A. P., and Vogelstein, B. (1983). *Anal. Biochem.* **132**, 6–13.
15. Church, G. M., and Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 1991–1995.
16. Fornace, A. J., Jr., Alamo, I., Jr., and Hollander, M. C. (1989) *Exp. Cell Res.* **182**, 61–74.
17. Chirgwin, J. J., Przbyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299.
18. Fornace, A. J., Jr., Alamo, I., Jr., Hollander, M. C., and Lamoreaux, E. (1989) *Nucleic Acids Res.* **17**, 1215–1230.
19. Britten, R. J., Graham, D. E., and Neufield, B. R. (1974) *in* Methods in Enzymology, (Grossman, L., and Moldave, K., Eds.), Vol. 29, pp. 363–417, Academic Press, New York.
20. Van Ness, J., and Hahn, W. E. (1982) *Nucleic Acids Res.* **10**, 8061–8077.
21. Gerard, G. J. (1985) *Focus* **7**, 7.
22. Hanahan, D. (1983) *J. Mol. Biol.* **166**, 557–580.